

Genegis: Computational Tools for Spatial Analyses of DNA Profiles with Associated Photo-Identification and Telemetry Records of Marine Mammals

C. Scott Baker
Oregon State University
Hatfield Marine Science Center
Newport, OR 97365-5296
phone: (541) 867-0255 fax: (541) 867-0138 email: scott.baker@oregonstate.edu

Dawn Wright
Department of Geosciences
Oregon State University
phone: (541) 737-1229 fax: (541) 737-1200 email: dawn@science.oregonstate.edu

John Calambokidis
Cascadia Research
218 ½ W. 4th Ave.
Olympia, WA 98501
phone: (360) 943-7325 ext 104 email: calambokidis@cascadiaresearch.org

Award Number: N000141110614
<http://mmi.oregonstate.edu/ccgl>

LONG-TERM GOALS

We are developing a suite of computational tools for improved visual exploration and spatial analysis of DNA profiles, with accompanying photo-identification records or telemetry tracks of marine mammals. Referred to as *geneGIS*, the tools provide the ability to display, browse, select, filter and summarize spatio-temporal relationships of these individual-based records and associated data from molecular markers and ecomarkers (e.g., stable isotopes). We have implemented *geneGIS* features on two platforms: 1) as programmatic enhancements of the web-based software, *Wildbook*, formerly the *Shepherd Project* (<http://www.wildme.org/wildbook/>); and, 2) as a toolbox in the desktop version of ArcGIS 10.1 or greater (<http://genegis.org/>). The software application and toolbox allow basic summaries of spatially selected data and export of data in standard tabular and database formats, as well as specialized formats required for programs commonly used in molecular ecology and capture-mark-recapture. The data export format complies with OBIS standards and the database architecture is compatible with the Arc Marine data model (Wright *et al.* 2007), providing a link with other datasets and tools needed for an integrated description of the genetic and environmental ‘seascape’ of cetaceans. The *Wildbook* implementation now includes online access to an integrated database of DNA profiles and photo-identification records derived from the SPLASH ocean-wide survey of humpback whales *Megaptera novaeangliae* in the North Pacific (Baker *et al.* in press, Barlow *et al.* 2011, Calambokidis *et al.* 2008). We have also implemented private *Wildbook* databases (password protected) for the following: 1) a subset of photo-identification sighting records and associated DNA

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 30 SEP 2013		2. REPORT TYPE		3. DATES COVERED 00-00-2013 to 00-00-2013	
4. TITLE AND SUBTITLE Genegis: Computational Tools for Spatial Analyses of DNA Profiles with Associated Photo-Identification and Telemetry Records of Marine Mammals				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Oregon State University, Hatfield Marine Science Center, 2030 SE Marine Science Drive, Hatfield, OR, 97365				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 16	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

profiles of right whales *Eubalaena glacialis* from the North Atlantic Right Whale Consortium; 2) DNA profiles of sperm whales *Physeter macrocephalus*, using samples collected during the 5-year *Voyage of the Odyssey*; and 3) DNA profiles of Hector's dolphins *Cephalorhynchus hectori* from Cloudy Bay, New Zealand, with funding from the US Marine Mammal Commission.

OBJECTIVES

The overall objectives can be ordered into five tasks (with related subtasks):

- Task 1: Develop database architecture following the Arc Marine data model for integration and display of DNA profiles with photo-identification and telemetry records in a stand-alone ArcGIS framework, and enhanced features of a web-based application (now referred to as *Wildbook*) initially designed for display and exploration of photo-identification catalogues.
- Task 2: Develop ArcGIS tools for data query, visual exploration and basic statistical summaries for spatial and temporal partitions of individual-based records (DNA profiles, photo-identification records and telemetry tracks).
- Task 3: Enhance user-directed spatial/temporal selection and export of individual-based records for advanced statistical analyses. This will include tools to export data compatible with existing software used for genetic analyses and capture-mark-recapture.
- Task 4: Demonstrate functionality of *geneGIS* and web-based application through importation and integration of existing large-scale datasets of photo-identification records from the Structure of Populations, Levels of Abundance and Status of Humpbacks program in the North Pacific (SPLASH) and associated DNA profiles from biopsy samples (*geneSPLASH*).
- Task 5: Prepare a comprehensive user guide for all software functions and analyses implemented in the system.

APPROACH

The computational developments follow two approaches: 1) tools and software enhancements within a web-based program for displaying individual identification photographs and information from linked DNA profiles, referred to as *geneGIS* in *Wildbook*; and 2) tools within ArcGIS, the most powerful widely-used software for GIS and advanced spatial analysis, referred to as *geneGIS* in ArcGIS. The intent is to benefit from the strengths of each approach while assuring compatibility and interoperability through a common database architecture and simplified import/export functions.

The web-based approach is being developed under subcontract to Jason Holmberg of *Wildbook* (formerly the *Shepherd Project*), with the support of John Calambokidis and Erin Falcone of Cascadia Research Collective. This approach takes advantage of an existing open-source software framework supporting capture-mark-recapture (CMR) studies of marine megafauna by the *Shepherd Project* (Holmberg *et al.* 2008). This software framework provides a scalable, web-based platform for CMR data management and was selected by Cascadia Research Collective to develop and host the SPLASH photo-identification catalog (on-line catalog created with the support of Pacific Life Foundation and available at <http://www.splashcatalog.org>). With support from the *geneGIS* initiative, the software framework of the *Shepherd Project* has been modified in *Wildbook* to include DNA profiles, providing a computational environment suitable for integrated studies of molecular ecology and CMR (<http://www.splashcatalog.org/latestgenegis/>).

The ArcGIS approach is being directed by the PI through Oregon State University(OSU), with support from Professor Dawn Wright (currently Esri Chief Scientist and adjunct professor at Oregon State University), Dori Dick, PhD candidate at OSU, Shaun Waldbridge, Esri, Ocean GIS Engineer, and members of the Marine Mammal Institute (MMI), including Tomas Follett and Debbie Steel. This approach takes advantage of previous experience with management of a whale telemetry database under the Arc Marine data model (Lord-Castillo *et al.* 2009; Wright *et al.* 2007). With support from the *geneGIS* initiative, we are developing tools to import and visualize spatial distributions and selection of individual identification records, as well as raster-based data extraction from environmental layers available in the ArcGIS environment.

WORK COMPLETED

- 1) Integrated database of photo-identification records and DNA profiles representing the *SPLASH* and *geneSPLASH* project;
- 2) Developed Import/Export functions for individual-based records into ArcGIS 10.1 (SRGD.csv and Arc Marine) and the *Wildbook*;
- 3) Developed *geneGIS* tool for spatial selection and comparison in ArcGIS 10.1;
- 4) Developed *geneGIS* tools for data extraction from environmental data layers in ArcGIS 10.1;
- 5) Implemented integrated *SPLASH/geneSPLASH* in web-accessible database through the *Wildbook*;
- 6) Developed *geneGIS* tools for custom analysis of molecular ecology and Capture-Mark-Recapture (CMR) in the the *Wildbook*;
- 7) Implementation and beta testing of other cetacean databases in *Wildbook* format (e.g., North Atlantic right whales, sperm whales and Hector's dolphins);
- 8) Supported continuing development of web-based User Manual and Developer Manual for the *Wildbook*;
- 9) Developed website and prepared draft installation manual for *geneGIS* tools in ArcGIS (<http://genegis.org/>); and,
- 10) Documented software development of toolbox in ArcGIS 10.1 through *Github* (<https://github.com/genegis/genegis>) and the *Wildbook* through the home page (<http://www.wildme.org/wildbook/>).

RESULTS

Integrated SPLASH/geneSPLASH databases

The *SPLASH* program provided a comprehensive dataset for implementation within ArcGIS and the *Wildbook*. At the inception of this project, *SPLASH* existed as a relational database in Microsoft Access with eight primary data tables containing effort, photo-identification, and tissue sampling records for humpback whales collected during five seasons of dedicated research effort in the North Pacific. This database and the associated photographic catalogue are maintained by Cascadia Research Collective. The photo-identification dataset was reconciled prior to this grant, to include over 18,400 encounters with 7,941 unique individuals and repeated encounters with individuals could be tracked throughout the database using an identifier known as the *SPLASHID* number (Calambokidis *et al.*

2008). A total of 5,675 tissue samples (mostly by biopsy darting) were also collected during SPLASH, of which about half were associated with a photo-identification encounter. From this total, 2,720 samples were selected for DNA profiling, including a sex marker, mitochondrial (mt) DNA haplotype sequencing and genotyping at 10 microsatellite loci. As expected, the 10 microsatellites were sufficiently variable to provide a second source of individual identification (Constantine *et al.* 2012; Madon *et al.* 2011). From the 2,720 samples, comparison of genotypes resolved 2,161 individuals. Referred to as geneSPLASH, the database of DNA profiles (i.e., a ‘DNA register’, DeSalle and Amato 2004) is maintained by the Cetacean Conservation and Genomics Laboratory, MMI, OSU.

The fully integrated database of photo-identification records and DNA profiles is one of the largest yet assembled for living whales and shows a complex pattern of migratory connections and population structure (Fig. 1). The results have been used to estimate the total abundance of humpback whales in the North Pacific, based on photo-identification (Barlow *et al.* 2011), and most recently, to describe the genetic structure of the population and propose delineation of Distinct Population Segments (Baker *et al.* in press).

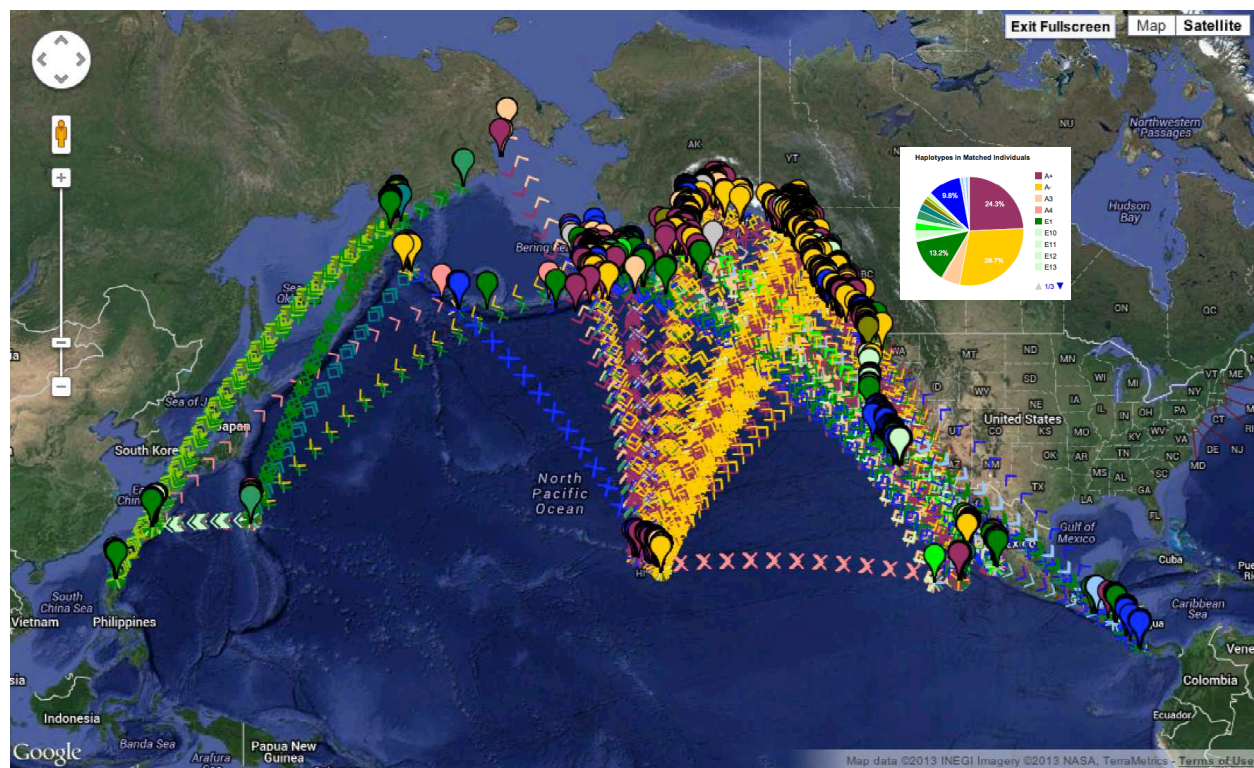


Figure 1: Use case for geneGIS in Wildbook showing the results of an Individual Search of the reconciled geneSPLASH databases for humpback whales. The results show the locations and migratory connections for $n = 1,981$ individuals with mtDNA haplotypes. The 28 different mtDNA haplotypes are color coded, showing the differentiation of regional wintering and feeding grounds. The pie chart shows the mtDNA haplotype frequencies of the 1,981 individuals. The chevrons indicate only a sequential link between encounters - they are not intended to represent a migratory path.

Shared database architecture and tabular input files

A database architecture has been agreed upon to accommodate relational data typical of those used in the collection of individual-based records from photo-identification and telemetry, with the associated collection of tissue samples for genetic analyses and ecomarkers. The architecture and nomenclature conform to Arc Marine and [Darwin Core](#) standards (Lord-Castillo *et al.* 2009; Wieczorek *et al.* 2012) where possible, and can accommodate the current databases developed for telemetry data and DNA profiles at MMI and SPLASH records at Cascadia Research Collective.

We have also developed import options for a simplified tabular structure (e.g., SRGD.csv or an Excel spreadsheet) similar to that more commonly used in molecular ecology and Capture-Mark-Recapture. The Spatially Referenced Genetic Data format (or SRGD) is a comma-separated value file providing for spatio-temporal records of encounters with individuals and associated DNA profiles. Additional data fields relating to the encounter or the individual, such as group size, behavioral roles or ecomarkers, can be placed after the primary fields. The intent is to allow an easy entry into ArcGIS and the *Wildbook* for access to the *geneGIS* tools by import of files that closely match the existing data formats used in genotyping and photo-identification.

geneGIS in *Wildbook*

Wildbook (<http://www.wildme.org/wildbook/>) is an open-source database framework designed to support studies in capture-mark-recapture (CMR), with recent enhancements to support molecular ecology and social ecology. The *Wildbook* is complementary to existing specialized programs for these studies, including Program MARK for estimates of population abundance and trends (White and Burnham 1999), Genepop for estimation of population genetic parameters (Rousset 2008), and SOCPROG for analysis of social affiliations (Whitehead 2009). The standard features of *Wildbook* now include:

- a scalable, collaborative platform for intelligent data storage and management, including advanced, consolidated searching;
- an easy-to-use suite of computation tools that can be extended to meet the needs of studies involving individual identification records and molecular ecology (e.g., photo-identification and DNA profiles);
- the export of data to specialized-analysis applications (e.g., Genepop, Program Capture), geographic software (e.g., Google Maps, ArcGIS); and standard biodiversity databases (e.g., GBIF and OBIS); and
- user management tools, including login and password protections for proprietary datasets.

The current version of this software and complete documentation of *Wildbook* capabilities are available at: <http://www.wildme.org/wildbook/>. *geneGIS* tools have been implemented through a number of software versions (now *Wildbook* 4.0) to include new capabilities for supporting molecular ecology. These enhancements have resulted in improved data search capabilities and visualization of molecular markers linked to individual identification, as displayed in the following use case.

Use Case: Spatial Selection and Regional Migratory Connections

The recent description of population structure of humpback whales in the North Pacific, based on samples collected in the SPLASH program (Baker *et al.* in press), showed a complex relationship between migratory destinations. Although significant genetic differences were found among both

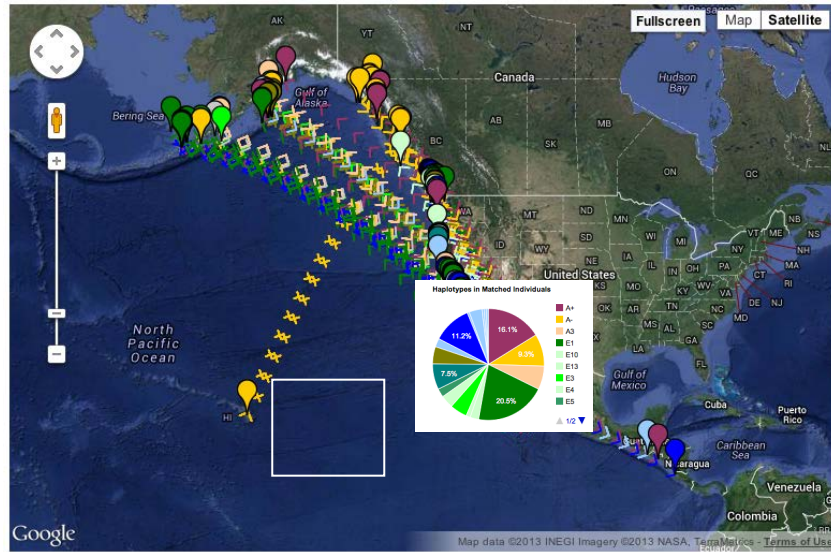
wintering grounds and feeding grounds, there was not a simple one-to-one relationship between seasonal habitats. This is effectively illustrated through an *Individual Search* function of the geneSPLASH database in *Wildbook* (Fig 2). Using the spatial tools to select the wintering grounds of Mainland and Baja, Mexico, returned $n = 159$ individuals with mtDNA haplotypes. The mapping function showed the range of resighting locations of these 159 individuals, based on both photo-IDs and genotypes. These migratory connections extend from the coast of California to the western Gulf of Alaska. The results also show an individual identified on an alternate winter ground (Hawaii), after migration to British Columbia. The spatial selection of the feeding grounds of California-Oregon returned $n = 129$ individuals with mtDNA haplotypes. The mapping function showed a much narrower range of migratory connections to Mainland and Baja, Mexico and Central America, with only a single connection to an alternate feeding ground in the western Gulf of Alaska.

The haplotype frequencies of California-Oregon clearly differ from those in Mainland and Baja, Mexico, but the complexity of the relationships between feeding and breeding grounds is only illuminated by the mapping of individual-based records, with associated genetic information. This has important implications for the designation of Distinct Population Segments under the US Endangered Species Act and for understanding Units to Conserve for ecosystem function (Baker *et al.* in press).

geneGIS in ArcGIS

Tool development in ArcGIS now includes data import, data export (to standard genetics programs) and custom analyses via a toolbox and a Python Add-in graphical user interface or GUI. The toolbox and installation instructions are available at <http://genegis.org/>.

a)



b)

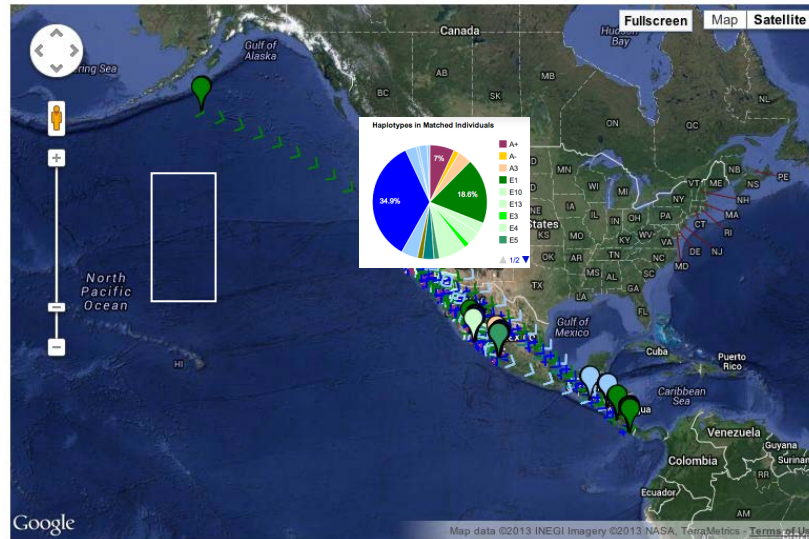


Figure 2. Use Case of geneGIS in Wildbook showing a spatially selected search of all individual humpback whales identified on a) Mainland and Baja Mexico wintering grounds, and b) the California-Oregon feeding grounds, during the SPLASH program, followed by a filter for individuals with mtDNA haplotypes. White boxes show initial spatial selection of individuals identified in a given region. The pie chart shows the mtDNA haplotype frequencies of individuals represented in that region. The chevrons indicate only a sequential link between encounters. They are not intended to represent a migratory path.

Table 1. The suite of tools developed, their function, and any relevant references available in geneGIS release 0.2 for ArcGIS 10.1.

Tool	Function
Import – Imports spatially reference data into ArcGIS	
1. Import	Creates a new file geodatabase (if one does not exist) and imports individual based records from the SRGD.csv input file into a feature class.
Export – Exports data from an ArcGIS feature class	
1. Export to Alleles in Space	Creates a file for use with Alleles in Space (AIS), software for the joint analysis of inter-individual spatial and genetic information. Miller 2005; http://www.marksgeneticssoftware.net/
2. Export to GenAlEx	Creates a text file formatted for use with GenAlEx, an MS Excel Add-in for population genetic analyses. Peakall and Smouse 2006, 2012; http://biology.anu.edu.au/GenAlEx
3. Export to Genepop	Creates a text file formatted for use with Genepop, a web based program for population genetic analyses. Raymond and Rousset 1995; Rousset 2008; http://genepop.curtin.edu.au
4. Export to KML	Creates a KML file, viewable in ArcGIS Explorer, ArcGlobe, and Google Earth.
5. Export to SPAGeDi	Creates a file formatted for SPAGeDi, software for the spatial pattern analysis of genetic diversity. Hardy and Vekemans 2002; http://ebe.ulb.ac.be/ebe/SPAGeDi.html
6. Export to SRGD	Creates a SRGD formatted table (CSV) for uploading to <i>Wildbook</i> , a software framework for mark-recapture studies. http://www.ecoceanusa.org/shepherd/doku.php
Custom Geodesic and Genetic Analyses	
1. Calculate F-statistics	Uses SPAGeDi software to calculate a variety of F-statistics and outputs them to a text file opening within ArcGIS. Hardy and Vekemans 2002; http://ebe.ulb.ac.be/ebe/SPAGeDi.html
1. Compute Distance Matrix	Computes a full pairwise geodesic distance matrix between all input locations. Calculations performed using Vincenty's formulae, accurate to within 0.5mm. Output is a CSV file.
2. Compute Distance Paths	Computes pairwise geodesic arcs connecting all input points. The arcs represent the shortest distance (great circle distance) between locations.
3. Individual Paths	Creates individual paths, linking a selected set of individuals across all locations they have been observed.
4. Extract Raster Values	Extracts values from one or more raster layers based on sample locations. Extracted point values are added to the attribute table of the feature class.
geneGIS Help	
1. Homepage	Takes the user to the project website homepage. http://genegis.org/
2. Documentation	Takes the user to the project documentation webpage. http://genegis.org/documentation.html

(Hardy and Vekemans 2002; Raymond and Rousset 1995; Rousset 2008; Peakall and Smouse 2006; Miller 2005)

Use Case: Individual Path analyses

A common analysis of individual identification records is the measurement of distances between serial sightings, within or between migratory habitats. Although ArcGIS has many options to measure distance between locations, calculating locations associated with a selected group of individuals is not available in the standard tools. The *geneGIS* tool *Individual Paths* creates individual paths, linking a selected set of individuals across all locations they have been encountered (see above, Table 1 and Fig. 3).

Issues arising – Software Licensing

One issue that has arisen during the development of the project is the appropriate licensing under which to release software. Although both developments of the ArcGIS toolbox and *Wildbook* are open source there are two different potential licenses, the Mozilla Public License, MPL (http://en.wikipedia.org/wiki/Mozilla_Public_License) and the GNU General Public License, GPL (http://en.wikipedia.org/wiki/GNU_General_Public_License). The *Wildbook* is currently licensed under GPL but the ArcGIS toolbox is currently unlicensed (although potentially downloadable through the website).

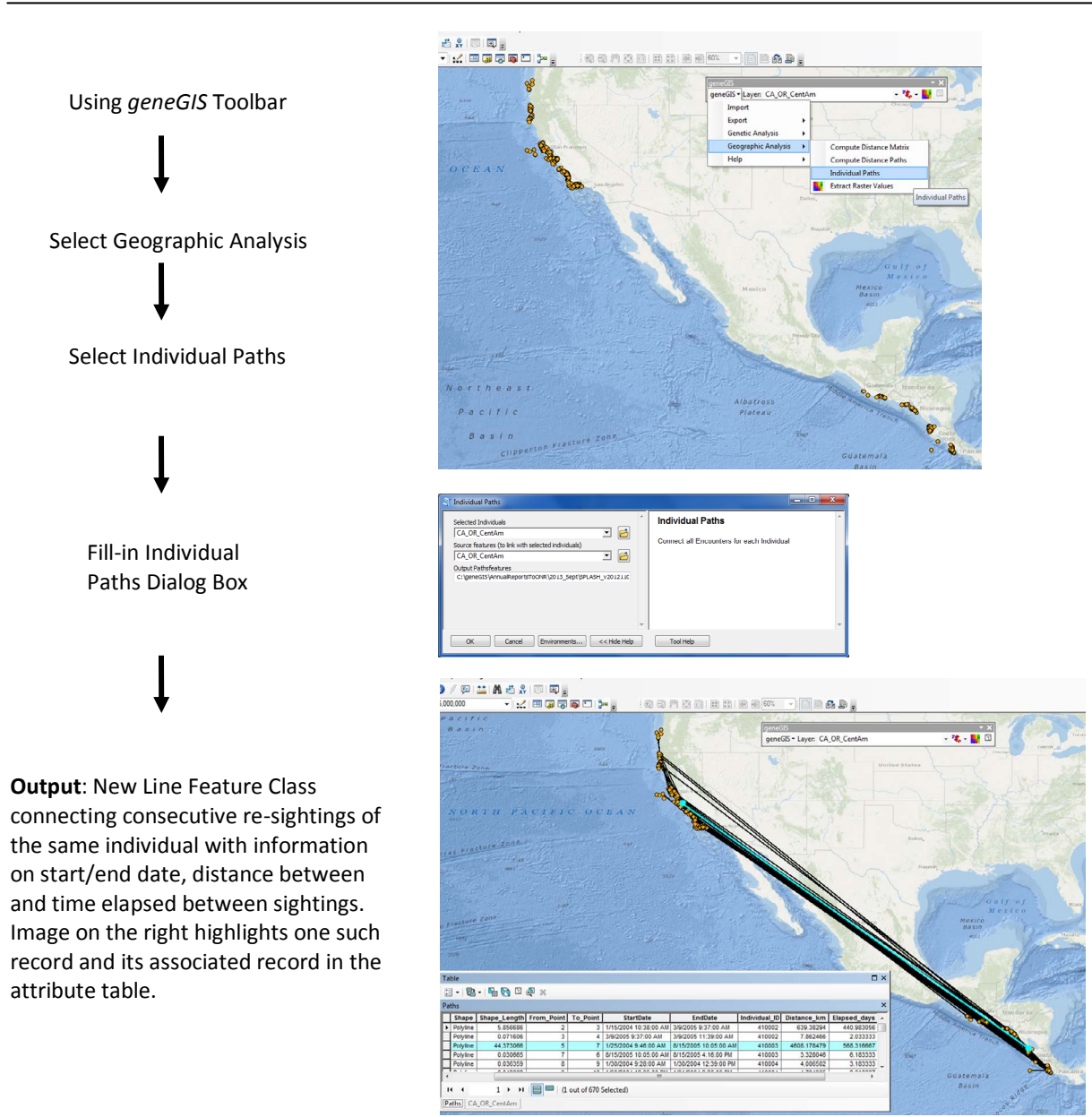


Figure 3: Use case of *geneGIS* in ArcGIS showing the spatial selection of individual encounters of humpback whales off California-Oregon and Central America. The Individual Paths tool then calculates all distances and times between serial sightings of an individual.

IMPACT/APPLICATIONS

With the development of *geneGIS* tools in *Wildbook* and ArcGIS, we expect to improve access to individual-based records and associated DNA profiles to the community of spatial modelers, and to contribute to the developing fields of landscape and seascape genetics (Etherington 2011; Vandergast *et al.* 2012). With the *geneGIS* enhancements of the *Wildbook* we are also developing a unified computational environment for Capture-Mark-Recapture and molecular ecology, including:

- distributed access to consolidated data and *geneGIS* analytic functions through a web browser, supporting distributed collaboration and the consolidation of multiple data sets;
- usable by a non-specialist, allowing non-GIS professionals to spatially explore and filter the data; and
- broader impact by contributing functionality to a open-source software framework also used by researchers for other species of marine megafauna (e.g., whale sharks and manta rays).

Now that *SPLASH/geneSPLASH* have been reconciled and implemented in both the Arc Marine model and the *Wildbook* framework, we have sought and obtained three additional datasets as exemplars for *geneGIS*: 1) a subset of photo-identification sighting records and associated DNA profiles from the North Atlantic Right Whale Consortium; 2) DNA profiles of sperm whales, using samples collected during the 5-year *Voyage of the Odyssey*; and 3) DNA profiles of Hector's dolphins from Cloudy Bay, New Zealand, with funding from the US Marine Mammal Commission. Each of these is currently considered proprietary and is accessible only through data-owner allocated passwords. More details included below under **Related Projects and Implementations**.

RELATED PROJECTS AND IMPLEMENTATIONS

Title: ‘*Examination of health effects and long-term impacts of deployments of multiple tag types on blue, humpback, and gray whales in the eastern North Pacific*’ with funding to Cascadia Research Collective, from the National Oceanographic Partnership Program (NOPP) and Interagency Committee on Ocean Science and Resource Management Integration (ICOSRMI). In collaboration with Cascadia Research Collective, the Marine Mammal Institute (MMI), Oregon State University (OSU) is assisting with the integration of photo-identification records and associated genetic samples, to improve understanding of long-term impact of satellite tagging. Preliminary results confirm the power of integrating genetic identity with photo-identification records to track the life history of individuals. By comparing DNA profiles from samples collected during satellite tagging, we have established a match with a whale tagged in Hawaii in 1999 and a whale sampled and photo-identified during the *SPLASH* program. The matching was conducted using the search features implemented in the *Wildbook* database of *SPLASH* project.

Title: ‘*Wildbook (formerly the Shepherd Project)*’. This project started as a collaborative software platform for globally coordinated whale shark research, which continues as <http://www.whaleshark.org/>. The success of this platform in managing and supporting the growth of the whale shark catalog led to its selection for the web-based implementation of the *SPLASH* Photo-ID Catalog (<http://www.splashcatalog.org>). Formerly know as the *Shepherd Project*, the ongoing software development was renamed *Wildbook* in mid 2013 to better reflect the intent to provide a computational environment in support of mark-recapture, molecular ecology, and social ecology studies based on individual records of vertebrate megafauna. Developments supported by the ONR through the *geneGIS* initiative are now acknowledged in recent applications to population structure and residency of whale sharks (Fox *et al.* 2013). We expect similar uptake of the *geneGIS* in *Wildbook* software by other researchers involved with studies of long-lived terrestrial or marine, species. Demonstration of the software at a recent workshop on advances in conservation genetics, hosted by University of Pretoria and the Southern African Wildlife College (<http://www.congen2013.co.za/>) resulted in interest from researchers studying African wild dogs, cheetahs and great white sharks.

Title: ‘*North Atlantic right whale Consortium*’. The North Atlantic right whale Consortium holds one of the most extensive catalogues of photo-ID records, with associated DNA profiles, available for any population of whales. We requested and were granted access to a subset of photo-ID sighting records (not the ID photographs themselves) and DNA profiles from the North Atlantic Right Whale (NARW) database through an application to the NARW Consortium. As with the SPLASH program, individual identification of NARWs has included both photo-identification and DNA profiling (Frasier *et al.* 2009), with photo-identification records maintained through the Digital Image Gathering and Information Tracking System (DIGITS) at the New England Aquarium. The dataset for 2001 and 2002 included $n = 6,553$ encounters of $n = 343$ individuals, of which 65 individuals had associated genetic information collected in those years (Frasier *et al.* 2007). The implementation was based on a small number of sub files provided independently by P. Hamilton and T. Frasier (in collaboration with B. White). The database is available at <http://www.splashcatalog.org/narw/> and password access was offered to the data providers for their review.

Title: ‘*Voyage of the Odyssey – genetic differentiation among equatorial populations of sperm whales*’, with funding from an International Fulbright Fellowship to PhD candidate, Alana Alexander. The *Voyage of the Odyssey* was a 5-year circumnavigation sponsored by the Ocean Alliance. The intent of the voyage was to document the health of the world’s oceans, in part by collecting biopsy samples from sperm whales for contaminant analyses (Wise Sr *et al.* 2009). Through a collaborative agreement with the Ocean Alliance, Alana Alexander was provided access to DNA from these samples for the purposes of describing the individual identity and population structure of the sampled whales. The DNA profiles of these samples have been uploaded for demonstration purposes to a proprietary *Wildbook* database (with password protection, Fig. 4). Permission was granted here to display the results of an Encounter Search of the database for reporting purposes. Results of these analyses are not in preparation for publication and intended for inclusion in Alana’s PhD thesis (e.g., (Alexander *et al.* 2013)).

Title: ‘*Hector’s dolphins in time and space*’ with funding from the US Marine Mammal Commission and Markam Fellowship support of PhD candidate, Rebecca Hamner. The New Zealand Threat Management Plan for the endemic Hector’s and Maui’s dolphins (subspecies of *Cephalorhynchus hectori*) lists the estimation of abundance as one of the four highest research priorities (New_Zealand_Department_of_Conservation_&_Ministry_of_Fisheries 2007). However, relatively small or subdivided populations of coastal dolphins pose particular problems for the estimation of abundance and trends. Here we attempted to improve the estimate for a local population of Hector’s dolphins in Cloudy Bay, New Zealand, using genotype capture-recapture by DNA profiling. This has previously proven powerful for estimating the abundance and subspecies identity of the Maui’s dolphin (Baker *et al.* 2012; Hamner *et al.* 2013b). During two field seasons, a total of 263 biopsy samples were collected, representing 148 individuals based on DNA profiling (Hamner *et al.* 2013a). The DNA profiles of these samples have been uploaded for demonstration purposes to a proprietary *Wildbook* database (with password protection), providing the opportunity to explore the *geneGIS* tools on a species with an annual range of movement on the order of 10s of kilometers, rather than the 1,000s of kilometers for migratory whales (Fig. 5).

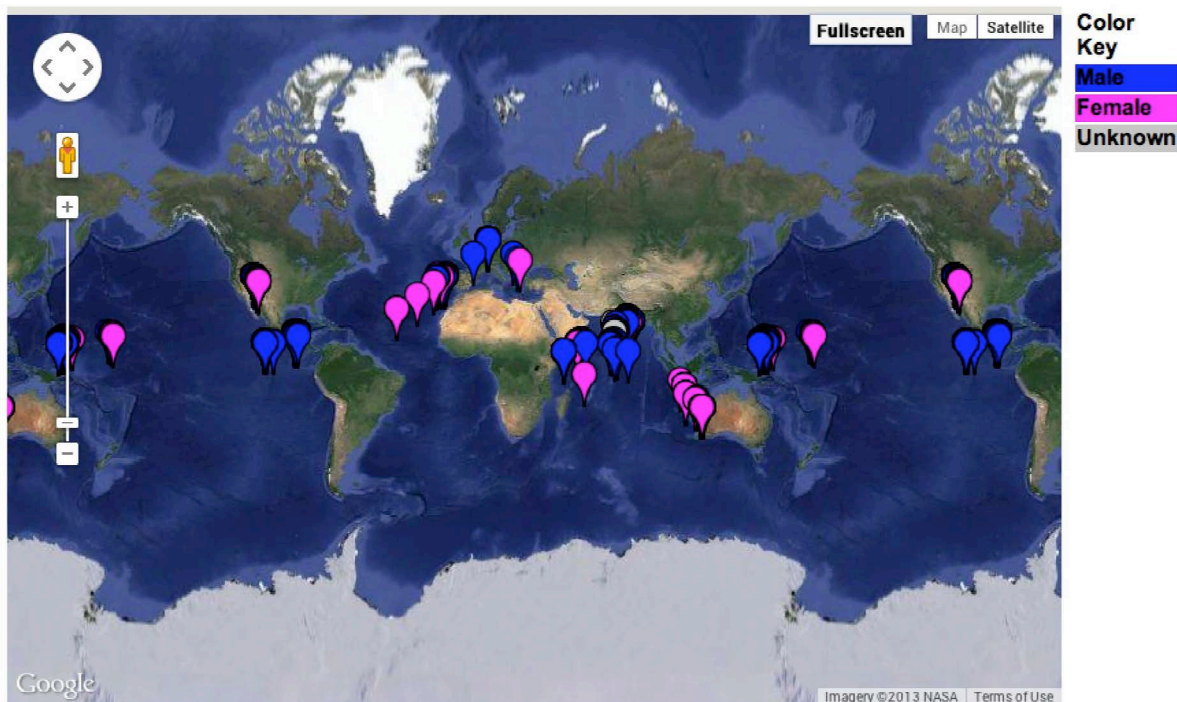


Figure 4: The location of 592 encounters with 512 individual sperm whales identified by DNA profiling from the collection during the *Voyage of the Odyssey*. The sex of the individuals encountered was identified by genetic markers.

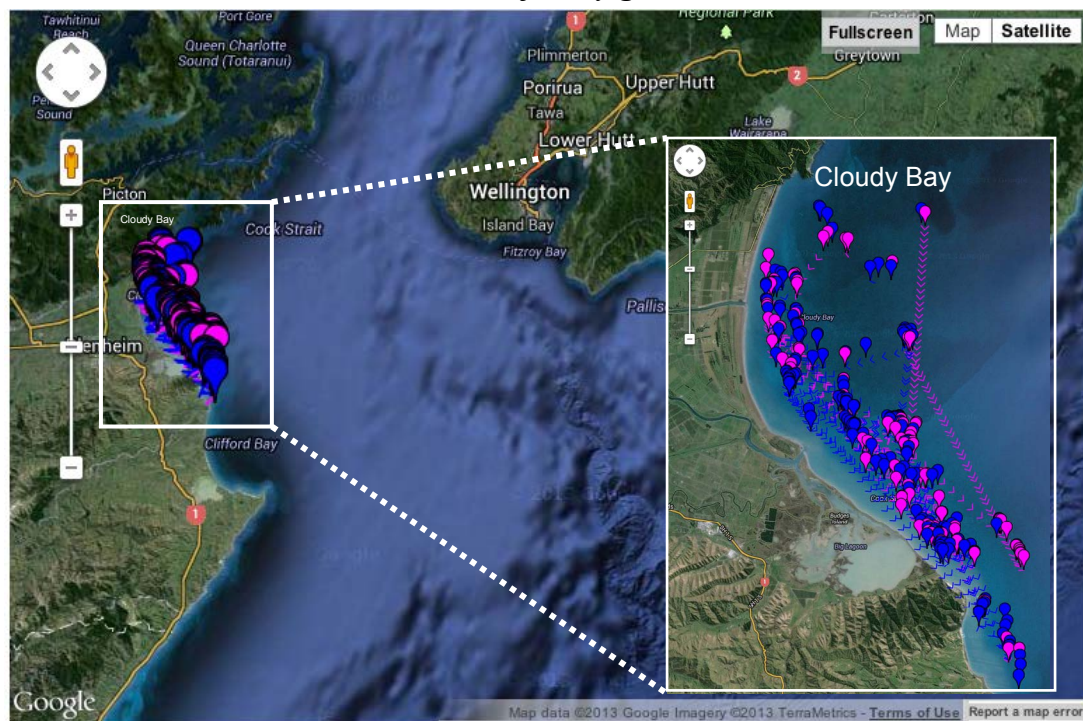


Figure 5: The location of 263 encounters with 148 individual Hector's dolphins identified by DNA profiling from the collection in Cloudy Bay, New Zealand, during the summers of 2011 and 2012. The sex of each individual (blue for male, pink for female) was identified by genetic markers. The inset shows an enlargement of Cloudy Bay to better show the local movement within and between the annual sampling periods.

Title: 'Humpback Whale (*Megaptera novaeangliae*) Genetic and Demographic Structure Within southeastern Alaska' with funding from a cooperative agreement with Glacier Bay National Park, in support of MSc student Sophie Pierszalowski. This implementation is still in progress.

CONFERENCE PRESENTATIONS

- Dick, D., S. Walbridge, D. Wright, J. Calambokidis, E. Falcone, D. Steel, T. Follett, J. Holmberg, C.S. Baker. 2013. *geneGIS*: computational tools for spatial analyses of DNA profiles with associated photo-identification records. Society for Marine Mammalogy, Dunedin 2013. Accepted for an oral presentation.
- Dick, D., D. Wright, J. Calambokidis, J. Holmberg, E. Falcone, T. Follett, D. Steel, B. Slikas, and C.S. Baker. 2012. *geneGIS*: computational tools for spatial analyses of DNA profiles associated and photo-identification records of whales and dolphins. Poster Presentation, The Wildlife Society 19th Annual Conference, Portland, OR, 13-18 October 2012.
- Holmberg, J. 2012. *The Shepherd Project*: A software framework for mark-recapture and molecular ecology. Oral Presentation, The Wildlife Society 19th Annual Conference, Portland, OR, 13-18 October 2012.

LITERATURE CITED

- Alexander, A., D. Steel, B. Slikas, K. Hoekzema, C. Carraher, M. Parks, R. Cronn and C.S. Baker. 2013. Low Diversity in the Mitogenome of Sperm Whales Revealed by Next-Generation Sequencing. *Genome Biology and Evolution* 5:113-129.
- Baker, C.S., R.M. Hamner, J. Cooke, D. Heimeier, M. Vant, D. Steel and R. Constantine. 2012. Low abundance and probable decline of the critically endangered Maui's dolphin estimated by genotype capture-recapture. *Animal Conservation* 16:224-233.
- Baker, C.S., D. Steel, J. Calambokidis, E.A. Falcone, U. Gozález-Peral, J. Barlow, A.M. Burdin, P.J. Clapham, J.K.B. Ford, C.M. Gabriele, U. Gozález-Peral, D. Mattila, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, J. Urbán-R., P. Wade, D. Weller, B.H. Witteveen and M. Yamaguchi. in press. Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology - Progress Series*.
- Barlow, J., J. Calambokidis, E.A. Falcone, C.S. Baker, A.M. Burdin, P.J. Clapham, J.K.B. Ford, C.M. Gabriele, R. LeDuc, D.K. Mattila, T.J. Quinn, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, J. Urbán R, P. Wade, D. Weller, B.H. Witteveen and M. Yamaguchi. 2011. Humpback whale abundance in the North Pacific estimated by photographic capture-recapture with bias correction from simulation studies. *Marine Mammal Science* 27:793-818.
- Calambokidis, J., E.A. Falcone, T.J. Quinn, A.M. Burdin, P.J. Clapham, J.K.B. Ford, C.M. Gabriele, R. LeDuc, D. Mattila, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, J.U. R., D. Weller, B.H. Witteveen, M. Yamaguchi, A. Bendlin, D. Camacho, K. Flynn, A. Havron, J. Huggins and N. Maloney. 2008. *SPLASH: Structure of Populations, Levels of Abundance and Status of Humpback Whales in the North Pacific*. U.S. Dept of Commerce, Western Administrative Center, Seattle, Washington.

- Constantine, R., J. Jackson, D. Steel, C.S. Baker, L. Brooks, D. Burns, P. Clapham, N. Hauser, B. Madon, D. Mattila, M. Oremus, M. Poole, J. Robbins, K. Thompson and C. Garrigue. 2012. Abundance of humpback whales in Oceania using photo-identification and microsatellite genotyping. *Marine Ecology Progress Series* 453:249-261.
- DeSalle, R. and G. Amato. 2004. The expansion of conservation genetics. *Nature Reviews Genetics* 5:702-712.
- Etherington, T.R. 2011. Python based GIS tools for landscape genetics: visualising genetic relatedness and measuring landscape connectivity. *Methods in Ecology and Evolution* 2:52-55.
- Fox, S., I. Foisy, R. De La Parra Venegas, B.E. Galván Pastoriza, R.T. Graham, E.R. Hoffmayer, J. Holmberg and S.J. Pierce. 2013. Population structure and residency of whale sharks *Rhincodon typus* at Utila, Bay Islands, Honduras. *Journal of Fish Biology*:n/a-n/a.
- Frasier, T.R., P.K. Hamilton, M.W. Brown, L.A. Conger, A.R. Knowlton, M.K. Marx, C.K. Slay, S.D. Kraus and B.N. White. 2007. Patterns of male reproductive success in a highly promiscuous whale species: the endangered North Atlantic right whale. *Molecular Ecology* 16:5277-5293.
- Frasier, T.R., P.K. Hamilton, M.W. Brown, S.D. Kraus and B.N. White. 2009. Sources and Rates of Errors in Methods of Individual Identification for North Atlantic Right Whales. *Journal of Mammalogy* 90:1246-1255.
- Hamner, R.M., R. Constantine, R. Mattlin, R. Waples and C.S. Baker. 2013a. Genotype capture-recapture estimates of abundance and effective population size of Hector's dolphins in Cloudy Bay, New Zealand. Report (SC/65/SM07) to the Scientific Committee of the International Whaling Commission.
- Hamner, R.M., R. Constantine, M. Oremus, M. Stanley, P. Brown and C. Scott Baker. 2013b. Long-range movement by Hector's dolphins provides potential genetic enhancement for critically endangered Maui's dolphin. *Marine Mammal Science*:n/a-n/a.
- Hardy, O.J. and X. Vekemans. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2:618-620.
- Holmberg, J., B. Norman and Z. Arzoumanian. 2008. Robust, comparable population metrics through collaborative photo-monitoring of whale sharks *Rhincodon typus* *Ecological Applications* 18:222-233.
- Lord-Castillo, B.K., B.R. Mate, D.J. Wright and T. Follett. 2009. A Customization of the Arc Marine Data Model to Support Whale Tracking via Satellite Telemetry. *Transactions in GIS* 13:63-83.
- Madon, B., O. Gimenez, B. McArdle, C. Scott Baker and C. Garrigue. 2011. A new method for estimating animal abundance with two sources of data in capture-recapture studies. *Methods in Ecology and Evolution*:no-no.
- Miller, M.P. 2005. Alleles In Space (AIS): Computer Software for the Joint Analysis of Interindividual Spatial and Genetic Information. *Journal of Heredity* 96:722-724.
- New_Zealand_Department_of_Conservation_&_Ministry_of_Fisheries. 2007. Hector's and Maui's Dolphin Threat Management Plan Draft for Public Consultation.<<http://www.fish.govt.nz/NR/rdonlyres/088EFD088C-E207-4798-9315-C4790AF4792FC4796FFB4792/4790/DRAFTTMPFINAL.pdf>>.

- Peakall, R. and P.E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288-229.
- Raymond, M. and F. Rousset. 1995. An exact test for population differentiation. *Evolution* 49:1280-1283.
- Rousset, F. 2008. Genepop, '007: a complete re-implementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources* 8:103-106.
- Vandergast, A.G., W.M. Perry, R.V. Lugo and S.A. Hathaway. 2012. Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. *Molecular Ecology Resources* 11:158-161.
- White, G.C. and K.P. Burnham. 1999. Program MARK: Survival estimation from populations of marked animals. *Bird Study* 46 (Supplement):120-138.
- Whitehead, H. 2009. SOCPROG programs: analysing animal social structures. *Behavioral Ecology and Sociobiology* 63:765-778.
- Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Döring, R. Giovanni, T. Robertson and D. Vieglais. 2012. Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS ONE* 7:e29715.
- Wise Sr, J.P., R. Payne, S.S. Wise, C. LaCerte, J. Wise, C. Gianios Jr, W. Douglas Thompson, C. Perkins, T. Zheng and C. Zhu. 2009. A global assessment of chromium pollution using sperm whales (*Physeter macrocephalus*) as an indicator species. *Chemosphere* 75:1461-1467.
- Wright, D.J., M.J. Blongewicz, P.N. Halpin and J. Breman. 2007. Arc Marine: GIS for a blue planet. Redlands, CA: ESRI Press.